(19) Europäisches Patentamt
European Patent Office
Office européen des brevets

(11) EP 3 952 437 A1

(12) EUROPEAN PATENT APPLICATION
published in accordance with Art. 153(4) EPC

(43) Date of publication:
09.02.2022 Bulletin 2022/06

(21) Application number: **20783725.3**

(22) Date of filing: **25.02.2020**

(51) International Patent Classification (IPC):
H04W 28/24 *(2009.01)*

(52) Cooperative Patent Classification (CPC):
H04W 28/24; H04L 41/5009; H04L 43/0858

(86) International application number:
PCT/CN2020/076623

(87) International publication number:
WO 2020/199792 (08.10.2020 Gazette 2020/41)

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR
Designated Extension States:
BA ME
Designated Validation States:
KH MA MD TN

(30) Priority: **02.04.2019 CN 201910261362**

(71) Applicant: **TENCENT TECHNOLOGY (SHENZHEN) COMPANY LIMITED**
Shenzhen, Guangdong 518057, (CN)

(72) Inventor: **WANG, Tao**
Shenzhen, Guangdong 518057 (CN)

(74) Representative: **Eisenführ Speiser**
Patentanwälte Rechtsanwälte PartGmbB
Postfach 10 60 78
28060 Bremen (DE)

(54) COMMUNICATION METHOD AND APPARATUS, COMPUTER-READABLE MEDIUM AND ELECTRONIC DEVICE

(57) Provided are a communication method and apparatus, a computer-readable medium and an electronic device. The communication method comprises: acquiring a total delay tolerance of uplink transmission and downlink transmission for a target service; generating a transmission delay indication message for the target service according to the total delay tolerance, wherein the transmission delay indication message comprises quality of service (QoS) flow indication information of the uplink transmission for the target service and QoS flow indication information of the downlink transmission for the target service; and configuring the transmission delay indication message for an access network entity, such that the access network entity monitors the transmission delay of the target service according to the transmission delay indication message. The technical solution of the embodiments of the present application can acquire corresponding total delay tolerances for different services, thereby realizing dynamic adjustment of the total delay tolerances for the services, and the technical solution can monitor the total delay of the uplink transmission and the downlink transmission of the target service, thereby being beneficial to meeting a QoS requirement of a service layer.

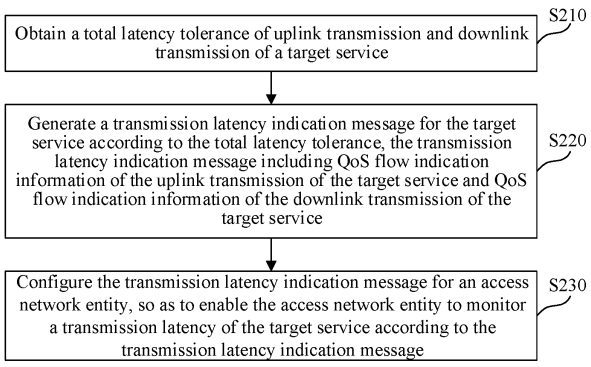

FIG. 2

EP 3 952 437 A1

Description

RELATED APPLICATION

[0001] This application claims priority to Chinese Patent Application No. 201910261362.9, entitled "COMMUNICATION METHOD AND APPARATUS, COMPUTER-READABLE MEDIUM AND ELECTRONIC DEVICE" and filed with the National Intellectual Property Administration, PRC on April 2, 2019, which is incorporated herein by reference in its entirety.

FIELD OF THE TECHNOLOGY

[0002] This application relates to the field of computer and communication technologies, and specifically, to a process of data communication.

BACKGROUND OF THE APPLICATION

[0003] To achieve higher rate experience, a greater bandwidth access capability, a lower latency, and highly reliable information exchange, the industry has proposed a network architecture based on a quality of service (QoS) flow. The QoS flow is the smallest granularity that provides QoS differentiation in a protocol data unit (PDU) session, and a QoS flow identifier (QFI) is used for identifying a QoS flow.

[0004] In addition, with the development of mobile communication technologies, a variety of new services have been derived, such as a cloud gaming service. The services have a relatively high requirement on a transmission latency as well as a requirement on a total latency of uplink (UL) and downlink (DL) transmission. However, the current QoS mechanism is too limited to meet the QoS requirements of the service layer.

SUMMARY

[0005] Embodiments of this application provide a communication method and apparatus, a computer-readable medium, and an electronic device, so as to monitor a total latency of UL transmission and DL transmission of a target service at least to some extent, to meet QoS requirements of a service layer.

[0006] Other features and advantages of this application become obvious through the following detailed descriptions, or may be partially learned through the practice of this application.

[0007] According to an aspect of the embodiments of this application, a communication method is provided, including: obtaining a total latency tolerance of UL transmission and DL transmission of a target service; generating a transmission latency indication message for the target service according to the total latency tolerance, the transmission latency indication message including QoS flow indication information of the UL transmission of the target service and QoS flow indication information of the DL transmission of the target service; and configuring the transmission latency indication message for an access network entity, so as to enable the access network entity to monitor a transmission latency of the target service according to the transmission latency indication message.

[0008] According to an aspect of the embodiments of this application, a communication method is provided, including: obtaining a transmission latency indication message configured by a core network (CN) entity for a target service, the transmission latency indication message being generated by the CN entity according to a total latency tolerance of UL transmission and DL transmission of the target service, and the transmission latency indication message including QoS flow indication information of the UL transmission of the target service and QoS flow indication information of the DL transmission of the target service; and monitoring a transmission latency of the target service according to the transmission latency indication message.

[0009] According to an aspect of the embodiments of this application, a communication apparatus is provided, including: an obtaining unit, configured to obtain a total latency tolerance of UL transmission and DL transmission of a target service; a generation unit, configured to generate a transmission latency indication message for the target service according to the total latency tolerance, the transmission latency indication message including QoS flow indication information of the UL transmission of the target service and QoS flow indication information of the DL transmission of the target service; and a configuration unit, configured to configure the transmission latency indication message for an access network entity, so as to enable the access network entity to monitor a transmission latency of the target service according to the transmission latency indication message.

[0010] In some embodiments of this application, based on the foregoing solutions, the obtaining unit is configured to: receive a notification message transmitted by an application function (AF) entity, and obtain the total latency tolerance according to the notification message; or obtain the total latency tolerance based on information of a contractual agreement with the AF entity; or obtain the total latency tolerance pre-configured by the AF entity.

[0011] In some embodiments of this application, based on the foregoing solutions, the generation unit is configured to: generate information indicating the total latency tolerance according to the total latency tolerance, and allocate QoS flow indication information respectively to the UL transmission and the DL transmission of the target service; and generate the transmission latency indication message according to the information indicating the total latency tolerance and the QoS flow indication information respectively allocated to the UL transmission and the DL transmission of the target service.

[0012] In some embodiments of this application, based on the foregoing solutions, the generation unit is config-

ured to: determine an UL latency tolerance and a DL latency tolerance of the target service according to the total latency tolerance; and generate the QoS flow indication information of the UL transmission according to the UL latency tolerance, and generate the QoS flow indication information of the DL transmission according to the DL latency tolerance.

[0013] In some embodiments of this application, based on the foregoing solutions, the generation unit is configured to: divide the total latency tolerance to obtain the UL latency tolerance and the DL latency tolerance of the target service; or receive a division status, notified by an AF entity, of the UL latency tolerance and the DL latency tolerance of the target service, and determine the UL latency tolerance and the DL latency tolerance according to the division status and the total latency tolerance; or receive priorities, notified by the AF entity, of the UL transmission and the DL transmission of the target service, and determine the UL latency tolerance and the DL latency tolerance according to the priorities and the total latency tolerance.

[0014] In some embodiments of this application, based on the foregoing solutions, the QoS flow indication information of the UL transmission of the target service includes: one piece of QoS flow indication information corresponding to all UL data of the target service; or QoS flow indication information respectively corresponding to different types of UL data of the target service, the different types of UL data being respectively corresponding to different QoS flow indication information.

[0015] In some embodiments of this application, based on the foregoing solutions, the QoS flow indication information of the DL transmission of the target service includes: one piece of QoS flow indication information corresponding to all DL data of the target service; or QoS flow indication information respectively corresponding to different types of DL data of the target service, the different types of DL data being respectively corresponding to different QoS flow indication information.

[0016] In some embodiments of this application, based on the foregoing solutions, when the communication method is performed by the CN entity, the obtaining unit is configured to: obtain the total latency tolerance from the AF entity by using a policy control function (PCF) entity; or obtain the total latency tolerance from the AF entity by using a network exposure function (NEF) entity and forward the total latency tolerance to the PCF entity.

[0017] According to an aspect of the embodiments of this application, a communication apparatus is provided, including: an obtaining unit, configured to obtain a transmission latency indication message configured by a CN entity for a target service, the transmission latency indication message being generated by the CN entity according to a total latency tolerance of UL transmission and DL transmission of the target service, and the transmission latency indication message including QoS flow indication information of the UL transmission of the target service and QoS flow indication information of the DL transmission of the target service; and a monitoring unit, configured to monitor a transmission latency of the target service according to the transmission latency indication message.

[0018] In some embodiments of this application, based on the foregoing solutions, the monitoring unit is configured to: monitor a total latency of the UL transmission and the DL transmission of the target service according to the total latency tolerance in a case that the transmission latency indication message further includes information indicating the total latency tolerance; monitor an UL transmission latency of the target service according to an UL latency tolerance in a case that the QoS flow indication information of the UL transmission includes the UL latency tolerance; and monitor a DL transmission latency of the target service according to a DL latency tolerance in a case that the QoS flow indication information of the DL transmission includes the DL latency tolerance.

[0019] In some embodiments of this application, based on the foregoing solutions, the QoS flow indication information of the UL transmission includes: QoS flow indication information respectively corresponding to different types of UL data of the target service, and the QoS flow indication information of the DL transmission includes QoS flow indication information respectively corresponding to different types of DL data of the target service; and the monitoring unit is configured to: monitor, according to the transmission latency indication message, a transmission latency of UL data of a specified type of the target service and a transmission latency of DL data of a specified type of the target service, the UL data of the specified type and the DL data of the specified type being located in the same PDU session or in different PDU sessions during transmission.

[0020] In some embodiments of this application, based on the foregoing solutions, the target service includes a cloud gaming service; and the UL data of the specified type includes game manipulation data, and the DL data of the specified type includes multimedia data obtained by rendering a game scene.

[0021] In some embodiments of this application, based on the foregoing solutions, the monitoring unit is configured to: monitor a transmission latency of the target service between user equipment (UE) and the access network entity according to the transmission latency indication message; and/or monitor a transmission latency of the target service between UE and a user plane function (UPF) entity according to the transmission latency indication message and latency information between the access network entity and the UPF entity.

[0022] According to an aspect of the embodiments of this application, a computer-readable medium is provided, storing a computer program, and the computer program, when executed by a processor, implementing the communication method according to the foregoing embodiments.

[0023] According to an aspect of the embodiments of

this application, an electronic device is provided, including: one or more processors; and a storage apparatus, configured to store one or more programs, and the one or more programs, when executed by the one or more processors, causing the one or more processors to implement the communication method according to the foregoing embodiments.

[0024] According to an aspect of the embodiments of this application, a computer program product including instructions is provided, the instructions, when run on a computer, causing the computer to perform the communication method in the foregoing embodiments.

[0025] In the technical solutions provided in some embodiments of this application, a total latency tolerance of UL transmission and DL transmission of a target service is obtained, a transmission latency indication message for the target service is generated according to the total latency tolerance, and the transmission latency indication message is configured for an access network entity. In this way, corresponding total latency tolerances can be obtained for different services, so that a total latency tolerance can be dynamically adjusted for a service. In addition, the access network entity can monitor the transmission latency of the target service according to the transmission latency indication message, which helps to monitor a total latency of the UL transmission and the DL transmission of the target service, thereby meeting QoS requirements of a service layer.

[0026] It is to be understood that the above general descriptions and the following detailed descriptions are merely for exemplary and explanatory purposes, and cannot limit this application.

BRIEF DESCRIPTION OF THE DRAWINGS

[0027] Accompanying drawings herein are incorporated into the specification and constitute a part of this specification, show embodiments that conform to this application, and are used for describing a principle of this application together with this specification. Apparently, the accompanying drawings described below are merely some embodiments of this application, and a person of ordinary skill in the art may further obtain other accompanying drawings according to the accompanying drawings without creative efforts. In the accompanying drawings:

FIG. 1 is a schematic diagram of an exemplary system architecture to which a technical solution according to an embodiment of this application is applicable.

FIG. 2 is a flowchart of a communication method according to an embodiment of this application.

FIG. 3 is a flowchart of generating a transmission latency indication message for a target service according to a total latency tolerance according to an embodiment of this application.

FIG. 4 is a flowchart of generating a transmission latency indication message for a target service according to a total latency tolerance according to an embodiment of this application.

FIG. 5 is a flowchart of a communication method according to an embodiment of this application.

FIG. 6 is a schematic diagram of a scenario of a cloud gaming service according to an embodiment of this application.

FIG. 7 is a schematic diagram of an interaction process among an AF entity, a CN entity, and a RAN entity according to an embodiment of this application.

FIG. 8 is a schematic diagram of an interaction process between an AF entity and a CN entity according to an embodiment of this application.

FIG. 9 is a schematic diagram of another interaction process between an AF entity and a CN entity according to an embodiment of this application.

FIG. 10 is a schematic diagram of another interaction process among an AF entity, a CN entity, and a RAN entity according to an embodiment of this application.

FIG. 11 is a schematic diagram of another interaction process among an AF entity, a CN entity, and a RAN entity according to an embodiment of this application.

FIG. 12 is a schematic diagram of a scenario of a cloud gaming service according to another embodiment of this application.

FIG. 13 is a block diagram of a communication apparatus according to an embodiment of this application.

FIG. 14 is a block diagram of a communication apparatus according to an embodiment of this application.

FIG. 15 is a schematic structural diagram of a computer system adapted to implement an electronic device according to an embodiment of this application.

DESCRIPTION OF EMBODIMENTS

[0028] At present, the exemplary implementations are described comprehensively with reference to the accompanying drawings. However, the examples of implemen-

tations may be implemented in a plurality of forms, and it is not to be understood as being limited to the examples described herein. Conversely, the implementations are provided to make this application more comprehensive and complete, and comprehensively convey the idea of the examples of the implementations to a person skilled in the art.

[0029] In addition, the described features, structures or characteristics may be combined in one or more embodiments in any appropriate manner. In the following descriptions, a lot of specific details are provided to give a comprehensive understanding of the embodiments of this application. However, a person of ordinary skill in the art is to be aware that, the technical solutions in this application may be implemented without one or more of the particular details, or another method, unit, apparatus, or step may be used. In other cases, well-known methods, apparatuses, implementations, or operations are not shown or described in detail, in order not to obscure the aspects of this application.

[0030] The block diagrams shown in the accompanying drawings are merely functional entities and do not necessarily correspond to physically independent entities. That is, the functional entities may be implemented in a software form, or in one or more hardware modules or integrated circuits, or in different networks and/or processor apparatuses and/or microcontroller apparatuses.

[0031] The flowcharts shown in the accompanying drawings are merely exemplary descriptions, do not need to include all content and operations/steps, and do not need to be performed in the described orders either. For example, some operations/steps may be further divided, while some operations/steps may be combined or partially combined. Therefore, an actual execution order may change according to an actual case.

[0032] FIG. 1 is a schematic diagram of an exemplary system architecture to which a technical solution according to an embodiment of this application is applicable.

[0033] As shown in FIG. 1, a system architecture 100 may include UE 101 (the UE 101 may be a smartphone shown in FIG. 1, or may be a tablet computer, a portable computer, a desktop computer, or the like), a base station 102, a CN entity 103, and an AF entity 104.

[0034] It is to be understood that FIG. 1 shows merely illustrative quantities of UEs 101, base stations 102, CN entities 103, and AF entities 104. There may be any quantities of UEs, base stations, CN entities, and AF entities according to an actual requirement.

[0035] In an embodiment of this application, the CN entity 103 may obtain a total latency tolerance, configured by the AF entity 104, of UL transmission and DL transmission of a target service. For example, the CN entity 103 may obtain the total latency tolerance by receiving a notification message transmitted by the AF entity 104. Alternatively, the CN entity 103 may obtain the total latency tolerance based on information of a contractual agreement with the AF entity 104. Alternatively, the CN entity 103 may obtain the total latency tolerance preconfigured by the AF entity 104.

[0036] In an embodiment of this application, after obtaining the total latency tolerance of the UL transmission and the DL transmission of the target service, the CN entity 103 may generate a transmission latency indication message for the target service according to the total latency tolerance, the transmission latency indication message including QoS flow indication information of the UL transmission of the target service and QoS flow indication information of the DL transmission of the target service, and then configure the transmission latency indication message for the base station 102, so as to enable the base station 102 to monitor a transmission latency of the target service according to the transmission latency indication message.

[0037] In an embodiment of this application, the base station 102 may monitor a transmission latency of the target service between the UE 101 and the base station 102 according to the transmission latency indication message; or monitor a transmission latency of the target service between the UE 101 and a UPF entity according to the transmission latency indication message and latency information between the base station 102 and the UPF entity.

[0038] The implementation details of the technical solutions in the embodiments of this application are described below in detail.

[0039] FIG. 2 is a flowchart of a communication method according to an embodiment of this application. The communication method may be performed by the CN entity 103 in FIG. 1. Referring to FIG. 2, the communication method includes at least S210 to S230. A detailed description is as follows:

S210: Obtain a total latency tolerance of UL transmission and DL transmission of a target service.

[0040] In an embodiment of this application, the total latency tolerance of the UL transmission and the DL transmission of the target service is obtained, so that total latency tolerances corresponding to different services can be obtained for the services, and a latency tolerance can be dynamically adjusted for a service. The target service may be a service that has a requirement on a total latency tolerance of UL transmission and DL transmission, for example, a cloud gaming service. The cloud gaming service is a manner of gaming based on cloud computing. In an operating mode of cloud gaming, UE needs to send game manipulation data (such as sensor data on the UE, action data detected by the UE, or the like) to a cloud end. The cloud end performs a rendering function according to the game manipulation data sent by the UE, and then transmits rendered multimedia data (including a game screen, sound data, or the like) to the UE via the network. The advantage of cloud gaming is that the UE does not need to perform complex rendering or computing, which reduces a hardware requirement on the UE. However, cloud gaming requires a network system to provide high-reliability and low-latency transmission for UL game manipulation data, and provide high-

bandwidth and low-latency transmission for DL multimedia data, and cloud gaming requires a total latency of UL transmission and DL transmission to be lower than a threshold.

[0041] In an embodiment of this application, the CN entity may receive a notification message transmitted by an AF entity, and then obtain the total latency tolerance of the UL transmission and the DL transmission of the target service according to the notification message.

[0042] In an embodiment of this application, the CN entity may obtain the total latency tolerance of the UL transmission and the DL transmission based on information of a contractual agreement with the AF entity. The information of the contractual agreement may be a service-level agreement (SLA).

[0043] In an embodiment of this application, the CN entity may obtain the total latency tolerance, pre-configured by the AF entity, of the UL transmission and the DL transmission.

[0044] In an embodiment of this application, the CN entity may include a PCF entity. The PCF entity may obtain the total latency tolerance of the UL transmission and the DL transmission from the AF entity. This case applies to a scenario where the PCF entity and the AF entity are in a trusted domain.

[0045] In an embodiment of this application, the CN entity may include a PCF entity and a NEF entity. The NEF entity may obtain the total latency tolerance of the UL transmission and the DL transmission from the AF entity, and forward the total latency tolerance to the PCF entity. This case applies to a scenario where the PCF entity and the AF entity are in different trusted domains.

[0046] Still referring to FIG. 2, S220: Generate a transmission latency indication message for the target service according to the total latency tolerance, the transmission latency indication message including QoS flow indication information of the UL transmission of the target service and QoS flow indication information of the DL transmission of the target service.

[0047] In an embodiment of this application, the QoS flow indication information may be a QFI. If applied to a 5G system, the QoS flow indication information may be a 5G QoS identifier (5QI).

[0048] In an embodiment of this application, the QoS flow indication information of the UL transmission of the target service may include: one piece of QoS flow indication information corresponding to all UL data of the target service. That is, in this embodiment, all UL data of the target service corresponds to one piece of QoS flow indication information. For example, UL data of the cloud gaming service includes game manipulation data and an acknowledgment of multimedia data sent by the cloud end, then the game manipulation data and the acknowledgment correspond to one piece of QoS flow indication information.

[0049] In an embodiment of this application, the QoS flow indication information of the UL transmission of the target service may include: QoS flow indication information respectively corresponding to different types of UL data of the target service, the different types of UL data being respectively corresponding to different QoS flow indication information. That is, in this embodiment, different types of UL data of the target service correspond to different QoS flow indication information. For example, UL data of the cloud gaming service includes game manipulation data and an acknowledgment of multimedia data sent by the cloud end, then the game manipulation data and the acknowledgment respectively correspond to different QoS flow indication information.

[0050] In an embodiment of this application, the QoS flow indication information of the DL transmission of the target service may include: one piece of QoS flow indication information corresponding to all DL data of the target service. That is, in this embodiment, all DL data of the target service corresponds to one piece of QoS flow indication information. For example, DL data of the cloud gaming service includes an acknowledgment of game manipulation data sent by the UE and multimedia data obtained after rendering, then the acknowledgment and the multimedia data correspond to one piece of QoS flow indication information.

[0051] In an embodiment of this application, the QoS flow indication information of the DL transmission of the target service may include: QoS flow indication information respectively corresponding to different types of DL data of the target service, the different types of DL data being respectively corresponding to different QoS flow indication information. That is, in this embodiment, different types of DL data of the target service correspond to different QoS flow indication information. For example, DL data of the cloud gaming service includes an acknowledgment of game manipulation data sent by the UE and multimedia data obtained after rendering, then the acknowledgment and the multimedia data respectively correspond to different QoS flow indication information.

[0052] In an embodiment of this application, as shown in FIG. 3, a process of generating a transmission latency indication message for the target service according to the total latency tolerance may include the following S310 and S320:

S310: Generate information indicating the total latency tolerance according to the total latency tolerance, and allocate QoS flow indication information respectively to the UL transmission and the DL transmission of the target service.

[0053] In an embodiment of this application, if the information indicating the total latency tolerance is generated, the QoS flow indication information respectively allocated to the UL transmission and the DL transmission of the target service may not require an absolute division of the total latency tolerance between the UL and the DL. That is, QoS flows respectively allocated to the UL transmission and the DL transmission of the target service may not clearly indicate a latency tolerance of the UL transmission and a latency tolerance of the DL transmission.

[0054] S320: Generate the transmission latency indication message according to the information indicating the total latency tolerance and the QoS flow indication information respectively allocated to the UL transmission and the DL transmission of the target service.

[0055] In an embodiment of this application, if the information indicating the total latency tolerance of the UL transmission and the DL transmission is generated, and the QoS flow indication information is allocated to the UL transmission and the DL transmission of the target service respectively, the transmission latency indication message may be generated according to the information, and the transmission latency indication message is then sent to an access network entity.

[0056] In the technical solution of the embodiment shown in FIG. 3, the total latency tolerance of the UL transmission and the DL transmission is indicated by using indication information other than the QoS flow indication information.

[0057] In an embodiment of this application, as shown in FIG. 4, a process of generating a transmission latency indication message for the target service according to the total latency tolerance may include the following S410 and S420:

S410: Determine an UL latency tolerance and a DL latency tolerance of the target service according to the total latency tolerance.

[0058] In an embodiment of this application, the CN entity (such as a PCF entity) may receive a division status, notified by an AF entity, of the UL latency tolerance and the DL latency tolerance of the target service, and determine the UL latency tolerance and the DL latency tolerance according to the division status and the total latency tolerance. That is, in the technical solution in this embodiment, the AF entity may directly notify a division status between the UL latency tolerance and the DL latency tolerance, and then the CN entity may directly determine the UL latency tolerance and the DL latency tolerance according to the division status.

[0059] In an embodiment of this application, the CN entity (such as a PCF entity) may receive priorities, notified by the AF entity, of the UL transmission and the DL transmission of the target service, and determine the UL latency tolerance and the DL latency tolerance according to the priorities and the total latency tolerance. That is, in the technical solution in this embodiment, the AF entity does not clearly notify a division status between the UL latency tolerance and the DL latency tolerance, but notifies priorities of the UL transmission and the DL transmission, and then the CN entity may divide the total latency tolerance according to the priorities to determine the UL latency tolerance and the DL latency tolerance. For example, if a priority of the UL transmission is higher than a priority of the DL transmission, the UL latency tolerance may be less than the DL latency tolerance, and then the total latency tolerance may be divided accordingly.

[0060] In an embodiment of this application, the CN entity (such as a PCF entity) may directly divide the total latency tolerance to obtain the UL latency tolerance and the DL latency tolerance of the target service. That is, in this embodiment, the CN entity may independently determine the UL latency tolerance and the DL latency tolerance according to the total latency tolerance. For example, the CN entity may determine latency requirements of different services on UL transmission and DL transmission according to historical data, and then independently determine a division status of an UL latency tolerance and a DL latency tolerance for a service accordingly.

[0061] S420: Generate the QoS flow indication information of the UL transmission according to the UL latency tolerance, and generate the QoS flow indication information of the DL transmission according to the DL latency tolerance.

[0062] In the technical solution of the embodiment shown in FIG. 4, the UL latency tolerance and the DL latency tolerance may be directly indicated by using the QoS flow indication information. Therefore, there is no need to additionally indicate the total latency tolerance of the UL transmission and the DL transmission.

[0063] Still referring to FIG. 2, S230: Configure the transmission latency indication message for an access network entity, so as to enable the access network entity to monitor a transmission latency of the target service according to the transmission latency indication message.

[0064] In an embodiment of this application, the access network entity may be a base station. The access network entity monitors a transmission latency of the target service according to the transmission latency indication message, that is, monitors an UL transmission latency and a DL transmission latency of the target service, to ensure that the UL transmission latency and the DL transmission latency do not exceed the total latency tolerance. The following embodiment illustrates the technical solutions of this application from the perspective of an access network entity.

[0065] FIG. 5 is a flowchart of a communication method according to an embodiment of this application. The communication method may be performed by an access network entity, and the access network entity may be the base station 102 shown in FIG. 1. Referring to FIG. 5, the communication method includes at least S510 and S520. A detailed description is as follows:

S510: Obtain a transmission latency indication message configured by a CN entity for a target service, the transmission latency indication message being generated by the CN entity according to a total latency tolerance of UL transmission and DL transmission of the target service, and the transmission latency indication message including QoS flow indication information of the UL transmission of the target service and QoS flow indication information of the DL transmission of the target service.

[0066] In an embodiment of this application, for a process in which the CN entity generates the transmission

latency indication message according to the total latency tolerance of the UL transmission and the DL transmission of the target service, reference may be made to the technical solutions of the foregoing embodiments.

[0067] S520: Monitor a transmission latency of the target service according to the transmission latency indication message.

[0068] In an embodiment of this application, in addition to the QoS flow indication information of the UL transmission of the target service and the QoS flow indication information of the DL transmission of the target service, the transmission latency indication message further includes information indicating the total latency tolerance of the UL transmission and the DL transmission, and then a total latency of the UL transmission and the DL transmission of the target service may be monitored according to the total latency tolerance.

[0069] In an embodiment of this application, an UL transmission latency of the target service may be monitored according to an UL latency tolerance in a case that the QoS flow indication information of the UL transmission includes the UL latency tolerance.

[0070] In an embodiment of this application, a DL transmission latency of the target service may be monitored according to a DL latency tolerance in a case that the QoS flow indication information of the DL transmission includes the DL latency tolerance.

[0071] In an embodiment of this application, if the access network entity detects that the total latency of the UL transmission and the DL transmission of the target service exceeds the total latency tolerance, or detects that the UL transmission latency of the target service exceeds the UL latency tolerance, or detects that the DL transmission latency of the target service exceeds the DL latency tolerance, a notification message may be sent to the CN entity or the AF entity, so that the CN entity or the AF entity can take corresponding measures.

[0072] In an embodiment of this application, the monitoring, by the access network entity, a transmission latency of the target service according to the transmission latency indication message may be monitoring a transmission latency of the target service between UE and the access network entity according to the transmission latency indication message.

[0073] In an embodiment of this application, the monitoring, by the access network entity, a transmission latency of the target service according to the transmission latency indication message may be monitoring a transmission latency of the target service between UE and a UPF entity according to the transmission latency indication message and latency information between the access network entity and the UPF entity. In this embodiment, because a latency between the access network entity and the UPF entity is generally fixed, the access network entity may monitor the transmission latency of the target service between the UE and the UPF entity according to the transmission latency of the target service between the UE and the access network entity and the latency between the access network entity and the UPF entity.

[0074] In an embodiment of this application, if the QoS flow indication information of the UL transmission includes QoS flow indication information respectively corresponding to different types of UL data of the target service, and the QoS flow indication information of the DL transmission includes QoS flow indication information respectively corresponding to different types of DL data of the target service, a transmission latency of UL data of a specified type of the target service and a transmission latency of DL data of a specified type of the target service may be monitored according to the transmission latency indication message. The UL data of the specified type and the DL data of the specified type are located in the same PDU session or in different PDU sessions during transmission.

[0075] For example, UL data of a cloud gaming service includes game manipulation data and an acknowledgment (referred to as a first acknowledgment, for the sake of distinction) of multimedia data sent by a cloud end, and DL data includes an acknowledgment (referred to as a second acknowledgment, for the sake of distinction) of game manipulation data sent by UE and multimedia data obtained after rendering. Then the game manipulation data, the first acknowledgment, the second acknowledgment, and the multimedia data respectively correspond to different QoS flow indication information. In this case, as the acknowledgment in the cloud gaming service is insignificant, it is possible to monitor only a transmission latency of the game manipulation data and a transmission latency of the multimedia data in the cloud gaming service. In addition, the game manipulation data and the second acknowledgment may be in one PDU session, and the multimedia data and the first acknowledgment may be in another PDU session; or the game manipulation data, the first acknowledgment, the second acknowledgment, and the multimedia data may be in one PDU session.

[0076] According to the technical solutions of the foregoing embodiments of this application, the total latency of the UL transmission and the DL transmission of the target service can be monitored, thereby meeting QoS requirements of a service layer.

[0077] The following describes the implementation details of the technical solutions of the embodiments of this application by using an example in which the target service is the cloud gaming service.

[0078] In an embodiment of this application, as shown in FIG. 6, in a scenario of a cloud gaming service, a cloud gaming application (APP) runs on UE 601, and an AF entity 602 may be a cloud rendering server, configured to perform rendering according to game manipulation data sent by the UE 601 to obtain multimedia data. UL data sent by the UE 601 includes A1 and B2, and DL data sent by the AF entity 602 includes A2 and B1. A1 represents game manipulation data, such as sensor data on the UE 601 and action data detected by the UE 601. B2

represents an ACK (acknowledgment) of multimedia data. A2 represents an ACK of the game manipulation data. B1 represents the multimedia data, such as a game screen and sound data.

[0079] The application scenario shown in FIG. 6 is a non-flow-splitting application scenario, that is, all UL data (including A1 and B2) corresponds to one QoS flow, and all DL data (including A2 and B1) corresponds to one QoS flow.

[0080] In the application scenario shown in FIG. 6, an interaction process among an AF entity, a CN entity, and a RAN entity (in this embodiment, description is made by using an example in which the RAN entity is a base station) is shown in FIG. 7, which may include the following steps:

S701: The AF entity notifies the CN entity of a total latency tolerance of UL+DL and a dynamic division of UL+DL.

[0081] In an embodiment of this application, the AF entity may notify the CN entity of the total latency tolerance of UL+DL and the dynamic division of UL+DL based on SLA interaction. Alternatively, a static configuration may be performed for the CN entity instead of notifying every time. Certainly, the AF entity may alternatively perform a dynamic configuration for the CN entity.

[0082] In an embodiment of this application, as shown in FIG. 8, the interaction process between the AF entity and the CN entity may be that the AF entity directly interacts with a PCF entity, then the PCF entity notifies an access and mobility management function (AMF) entity of a determined latency tolerance, and the AMF entity configures the latency tolerance for the base station. The technical solution of this embodiment is applicable to an application scenario where the AF entity and the PCF entity are in the same trusted domain, such as a scenario where a network operator itself deploys the AF entity.

[0083] In an embodiment of this application, as shown in FIG. 9, the interaction process between the AF entity and the CN entity may be that the AF entity interacts with a PCF entity by using a NEF entity, then the PCF entity notifies an AMF entity of a determined latency tolerance, and the AMF entity configures the latency tolerance for the base station. The technical solution of this embodiment is applicable to an application scenario where the AF entity and the PCF entity are in different trusted domains, such as a scenario where a third party deploys the AF entity.

[0084] In an embodiment of this application, the AF entity may further adjust a bitrate for sending the multimedia data to the UE according to a QoS monitoring result, a predicted network condition, and a network resource condition that are provided by a network system (such as a 5G system). When the QoS monitoring result is good, the predicted network condition is good, or network resources are sufficient, the bitrate for sending the multimedia data to the UE can be increased so as to improve the definition of a cloud gaming screen displayed on the UE while ensuring fluency. When the QoS monitoring result is poor, the predicted network condition is poor, or network resources are insufficient, the bitrate for sending the multimedia data to the UE is decreased so as to improve fluency of a cloud gaming screen displayed on the UE.

[0085] S702: The CN entity determines UL and DL latency tolerances according to the total latency tolerance of UL+DL and the dynamic division of UL+DL.

[0086] In an embodiment of this application, because the AF entity has notified the total latency tolerance of UL+DL and the dynamic division of UL+DL, the CN entity can determine the UL and DL latency tolerances accordingly, and then allocate corresponding 5QI values for UL data and DL data according to the UL and DL latency tolerances. In this case, QoS flows allocated respectively for the UL data and the DL data can indicate a latency tolerance of UL transmission and a latency tolerance of DL transmission.

[0087] S703: The CN entity configures the latency tolerances for the base station so as to enable the base station to monitor a latency dynamically.

[0088] In an embodiment of this application, the CN entity may configure, for the base station, the 5QI values respectively allocated for the UL data and the DL data. During UL and DL scheduling, the base station may monitor whether an UL transmission latency exceeds the latency tolerance of the UL transmission, and monitor whether a DL transmission latency exceeds the latency tolerance of the DL transmission. If the UL transmission latency exceeds the latency tolerance of the UL transmission, or the DL transmission latency exceeds the latency tolerance of the DL transmission, the base station may notify the CN entity and the AF entity.

[0089] In the application scenario shown in FIG. 6, another interaction process among an AF entity, a CN entity, and a RAN entity is shown in FIG. 10, which may include the following steps:

S1001: The AF entity notifies the CN entity of a total latency tolerance of UL+DL, and UL and DL transmission priorities.

[0090] The interaction manner between the AF entity and the CN entity in this embodiment is the same as that in the foregoing embodiment, and is not repeated herein.

[0091] S1002: The CN entity determines UL and DL latency tolerances according to the total latency tolerance of UL+DL and the UL and DL transmission priorities.

[0092] In an embodiment of this application, because the AF entity notifies of the total latency tolerance of UL+DL and the UL and DL transmission priorities, the CN entity may determine the UL and DL latency tolerances accordingly. For example, if a priority of UL transmission is higher than a priority of DL transmission, the UL latency tolerance may be less than the DL latency tolerance, and then the total latency tolerance can be divided accordingly. Then the CN entity allocates corresponding 5QI values for UL data and DL data according to the UL and DL latency tolerances. In this case, QoS flows allocated respectively for the UL data and the DL data can indicate the latency tolerance of the UL trans-

mission and the latency tolerance of the DL transmission.

[0093] S1003: The CN entity configures the latency tolerances for the base station so as to enable the base station to monitor a latency dynamically.

[0094] In an embodiment of this application, the CN entity may configure, for the base station, the 5QI values respectively allocated for the UL data and the DL data. During UL and DL scheduling, the base station may monitor whether an UL transmission latency exceeds the latency tolerance of the UL transmission, and monitor whether a DL transmission latency exceeds the latency tolerance of the DL transmission. If the UL transmission latency exceeds the latency tolerance of the UL transmission, or the DL transmission latency exceeds the latency tolerance of the DL transmission, the base station may notify the CN entity and the AF entity.

[0095] In the application scenario shown in FIG. 6, another interaction process among an AF entity, a CN entity, and a RAN entity is shown in FIG. 11, which may include the following steps:

S1101: The AF entity notifies the CN entity of a total latency tolerance of UL+DL.

[0096] The interaction manner between the AF entity and the CN entity in this embodiment is the same as the foregoing embodiment, and is not repeated herein.

[0097] S1102: The CN entity allocates 5QI values to the UL and the DL according to the total latency tolerance of UL+DL.

[0098] In an embodiment of this application, when allocating the 5QI values to the UL and the DL according to the total latency tolerance of UL+DL, the CN entity may first determine UL and DL latency tolerances. For example, the CN entity may determine a latency requirement of the cloud gaming service on UL transmission and DL transmission based on historical data, then independently determine a division status of an UL latency tolerance and a DL latency tolerance for the cloud gaming service accordingly, divide the total latency tolerance to obtain the UL and DL latency tolerances, and allocate the 5QI values to the UL and the DL accordingly. In this case, QoS flows allocated respectively for UL data and DL data can indicate the latency tolerance of the UL transmission and the latency tolerance of the DL transmission.

[0099] In an embodiment of this application, when the CN entity allocates the 5QI values to the UL and the DL according to the total latency tolerance of UL+DL, an absolute division of the latency tolerance between the UL and the DL may not be required. In this case, the CN entity may introduce a new indication message to indicate the total latency tolerance of UL+DL to the base station.

[0100] S1103: The CN entity configures the latency tolerances for the base station so as to enable the base station to monitor a latency dynamically.

[0101] In an embodiment of this application, if QoS flows allocated by the CN entity respectively for the UL data and the DL data can indicate the latency tolerance of the UL transmission and the latency tolerance of the DL transmission, the CN entity may configure, for the base station, the 5QI values respectively allocated for the UL data and the DL data. Then during UL and DL scheduling, the base station may monitor whether an UL transmission latency exceeds the latency tolerance of the UL transmission, and monitor whether a DL transmission latency exceeds the latency tolerance of the DL transmission. If the UL transmission latency exceeds the latency tolerance of the UL transmission, or the DL transmission latency exceeds the latency tolerance of the DL transmission, the base station may notify the CN entity and the AF entity.

[0102] In an embodiment of this application, if QoS flows allocated by the CN entity respectively for the UL data and the DL data do not require an absolute division of the latency tolerance between the UL and the DL, and the CN entity introduces a new indication message to indicate the total latency tolerance of UL+DL to the base station, then during UL and DL scheduling, the base station may monitor whether a total latency of the UL transmission and the DL transmission exceeds the total latency tolerance. If the total latency of the UL transmission and the DL transmission exceeds the total latency tolerance, the base station may notify the CN entity and the AF entity.

[0103] In another embodiment of this application, as shown in FIG. 12, in a scenario of a cloud gaming service, a cloud gaming APP runs on UE 1201, and an AF entity 1202 may be a cloud rendering server, configured to perform rendering according to game manipulation data sent by the UE 1201 to obtain multimedia data. UL data sent by the UE 1201 includes A1 and B2, and DL data sent by the AF entity 1202 includes A2 and B1. A1 represents game manipulation data, such as sensor data on the UE 1201 and action data detected by the UE 1201. B2 represents an ACK (acknowledgment) of multimedia data. A2 represents an ACK of the game manipulation data. B1 represents the multimedia data, such as a game screen and sound data.

[0104] The application scenario shown in FIG. 12 is a flow-splitting application scenario, that is, the UL data A1 and B2 correspond to different QoS flows, and the DL data A2 and B1 also correspond to different QoS flows. A1 and A2 may be in one PDU session, and B1 and B2 may be in one PDU session; or A1, A2, B1, and B2 may be in one PDU session.

[0105] In the application scenario shown in FIG. 12, an interaction process among an AF entity, a CN entity, and a RAN entity is similar to the examples shown in FIG. 7 to FIG. 11 in the foregoing embodiments, and differences are specifically as follows:

(1) The CN entity needs to allocate QoS flows for different UL data (A1 and B2) and different DL data (A2 and B1) respectively when allocating QoS flows for UL data and DL data respectively.

(2) When the base station performs UL and DL

scheduling, if the QoS flows can indicate a latency tolerance of UL transmission and a latency tolerance of DL transmission, the base station may monitor only whether an UL transmission latency of A1 (the game manipulation data) exceeds a corresponding latency tolerance of the UL transmission and whether a DL transmission latency of B1 (the multimedia data) exceeds a corresponding latency tolerance of the DL transmission. If the UL transmission latency of A1 exceeds the corresponding latency tolerance of the UL transmission, or the DL transmission latency of B1 exceeds the corresponding latency tolerance of the DL transmission, the base station may notify the CN entity and the AF entity.

(3) When the base station performs UL and DL scheduling, if QoS flows allocated by the CN entity for the UL data and the DL data respectively do not require an absolute division of the latency tolerance between the UL and the DL, and the CN entity introduces a new indication message to indicate a total latency tolerance of UL+DL to the base station, the base station may monitor whether a sum of a UL transmission latency of A1 (the game manipulation data) and a DL transmission latency of B1 (the multimedia data) exceeds the total latency tolerance. If the sum of the UL transmission latency of A1 and the DL transmission latency of B1 exceeds the total latency tolerance, the base station may notify the CN entity and the AF entity.

[0106] It can be learned that in the flow-splitting application scenario shown in FIG. 12, the technical solution of this embodiment of this application can ensure a total latency tolerance of some data flows in the UL transmission and some data flows in the DL transmission, which effectively improves the flexibility of a QoS flow mechanism.

[0107] According to the technical solutions of the embodiments shown in FIG. 6 to FIG. 12, not only the total latency of the UL transmission and the DL transmission of the cloud gaming service is monitored, but also the dynamic latency tolerances of the UL and the DL are implemented, which effectively meets QoS requirements of a service layer. The technical solutions of the foregoing embodiments of this application are also applicable to other services that need to ensure a total latency of UL transmission and DL transmission.

[0108] The following describes apparatus embodiments of this application, and the apparatus embodiments may be used for performing the communication methods in the foregoing embodiments of this application. For details not disclosed in the apparatus embodiments of this application, reference may be made to the foregoing communication method embodiments of this application.

[0109] FIG. 13 is a block diagram of a communication apparatus according to an embodiment of this application.

[0110] Referring to FIG. 13, a communication apparatus 1300 according to an embodiment of this application includes: an obtaining unit 1302, a generation unit 1304, and a configuration unit 1306.

[0111] The obtaining unit 1302 is configured to obtain a total latency tolerance of UL transmission and DL transmission of a target service. The generation unit 1304 is configured to generate a transmission latency indication message for the target service according to the total latency tolerance, the transmission latency indication message including QoS flow indication information of the UL transmission of the target service and QoS flow indication information of the DL transmission of the target service. The configuration unit 1306 is configured to configure the transmission latency indication message for an access network entity, so as to enable the access network entity to monitor a transmission latency of the target service according to the transmission latency indication message.

[0112] In some embodiments of this application, based on the foregoing solutions, the obtaining unit 1302 is configured to: receive a notification message transmitted by an AF entity, and obtain the total latency tolerance according to the notification message; or obtain the total latency tolerance based on information of a contractual agreement with the AF entity; or obtain the total latency tolerance pre-configured by the AF entity.

[0113] In some embodiments of this application, based on the foregoing solutions, the generation unit 1304 is configured to: generate information indicating the total latency tolerance according to the total latency tolerance, and allocate QoS flow indication information respectively to the UL transmission and the DL transmission of the target service; and generate the transmission latency indication message according to the information indicating the total latency tolerance and the QoS flow indication information respectively allocated to the UL transmission and the DL transmission of the target service.

[0114] In some embodiments of this application, based on the foregoing solutions, the generation unit 1304 is configured to: determine an UL latency tolerance and a DL latency tolerance of the target service according to the total latency tolerance; and generate the QoS flow indication information of the UL transmission according to the UL latency tolerance, and generate the QoS flow indication information of the DL transmission according to the DL latency tolerance.

[0115] In some embodiments of this application, based on the foregoing solutions, the generation unit 1304 is configured to: divide the total latency tolerance to obtain the UL latency tolerance and the DL latency tolerance of the target service; or receive a division status, notified by an AF entity, of the UL latency tolerance and the DL latency tolerance of the target service, and determine the UL latency tolerance and the DL latency tolerance according to the division status and the total latency tolerance; or receive priorities, notified by the AF entity, of the

UL transmission and the DL transmission of the target service, and determine the UL latency tolerance and the DL latency tolerance according to the priorities and the total latency tolerance.

[0116] In some embodiments of this application, based on the foregoing solutions, the QoS flow indication information of the UL transmission of the target service includes: one piece of QoS flow indication information corresponding to all UL data of the target service; or QoS flow indication information respectively corresponding to different types of UL data of the target service, the different types of UL data being respectively corresponding to different QoS flow indication information.

[0117] In some embodiments of this application, based on the foregoing solutions, the QoS flow indication information of the DL transmission of the target service includes: one piece of QoS flow indication information corresponding to all DL data of the target service; or QoS flow indication information respectively corresponding to different types of DL data of the target service, the different types of DL data being respectively corresponding to different QoS flow indication information.

[0118] In some embodiments of this application, based on the foregoing solutions, when the communication method is performed by the CN entity, the obtaining unit 1302 is configured to: obtain the total latency tolerance from the AF entity by using a PCF entity; or obtain the total latency tolerance from the AF entity by using a NEF entity and forward the total latency tolerance to the PCF entity.

[0119] FIG. 14 is a block diagram of a communication apparatus according to an embodiment of this application.

[0120] Referring to FIG. 14, a communication apparatus 1400 according to an embodiment of this application includes: an obtaining unit 1402 and a monitoring unit 1404.

[0121] The obtaining unit 1402 is configured to obtain a transmission latency indication message configured by a CN entity for a target service, the transmission latency indication message being generated by the CN entity according to a total latency tolerance of UL transmission and DL transmission of the target service, and the transmission latency indication message including QoS flow indication information of the UL transmission of the target service and QoS flow indication information of the DL transmission of the target service. The monitoring unit 1404 is configured to monitor a transmission latency of the target service according to the transmission latency indication message.

[0122] In some embodiments of this application, based on the foregoing solutions, the monitoring unit 1404 is configured to: monitor a total latency of the UL transmission and the DL transmission of the target service according to the total latency tolerance in a case that the transmission latency indication message further includes information indicating the total latency tolerance; monitor an UL transmission latency of the target service according to an UL latency tolerance in a case that the QoS flow indication information of the UL transmission includes the UL latency tolerance; and monitor a DL transmission latency of the target service according to a DL latency tolerance in a case that the QoS flow indication information of the DL transmission includes the DL latency tolerance.

[0123] In some embodiments of this application, based on the foregoing solutions, the QoS flow indication information of the UL transmission includes: QoS flow indication information respectively corresponding to different types of UL data of the target service, and the QoS flow indication information of the DL transmission includes QoS flow indication information respectively corresponding to different types of DL data of the target service; and the monitoring unit 1404 is configured to: monitor, according to the transmission latency indication message, a transmission latency of UL data of a specified type of the target service and a transmission latency of DL data of a specified type of the target service, the UL data of the specified type and the DL data of the specified type being located in the same PDU session or in different PDU sessions during transmission.

[0124] In some embodiments of this application, based on the foregoing solutions, the target service includes a cloud gaming service; and the UL data of the specified type includes game manipulation data, and the DL data of the specified type includes multimedia data obtained by rendering a game scene.

[0125] In some embodiments of this application, based on the foregoing solutions, the monitoring unit 1404 is configured to: monitor a transmission latency of the target service between UE and the access network entity according to the transmission latency indication message; and/or monitor a transmission latency of the target service between UE and a UPF entity according to the transmission latency indication message and latency information between the access network entity and the UPF entity.

[0126] FIG. 15 is a schematic structural diagram of a computer system adapted to implement an electronic device according to an embodiment of this application.

[0127] The computer system 1500 of the electronic device shown in FIG. 15 is merely an example, and does not constitute any limitation on functions and use ranges of the embodiments of this application.

[0128] As shown in FIG. 15, the computer system 1500 includes a central processing unit (CPU) 1501, which may perform various suitable actions and processing based on a program stored in a read-only memory (ROM) 1502 or a program loaded from a storage part 1508 into a random access memory (RAM) 1503, for example, perform the method described in the foregoing embodiments. The RAM 1503 further stores various programs and data required for system operations. The CPU 1501, the ROM 1502, and the RAM 1503 are connected to each other through a bus 1504. An input/output (I/O) interface 1505 is also connected to the bus 1504.

[0129] The following components are connected to the I/O interface 1505: an input part 1506 including a keyboard, a mouse, or the like, an output part 1507 including a cathode ray tube (CRT), a liquid crystal display (LCD), a speaker, or the like, a storage part 1508 including a hard disk, or the like, and a communication part 1509 including a network interface card such as a local area network (LAN) card or a modem. The communication part 1509 performs communication processing via a network such as the Internet. A driver 1510 is also connected to the I/O interface 1505 as required. A removable medium 1511, such as a magnetic disk, an optical disc, a magneto-optical disk, or a semiconductor memory, is installed on the driver 1510 as required, so that a computer program read from the removable medium is installed into the storage part 1508 as required.

[0130] Particularly, according to an embodiment of this application, the processes described in the following by referring to the flowcharts may be implemented as computer software programs. For example, this embodiment of this application includes a computer program product, the computer program product includes a computer program carried on a computer-readable medium, and the computer program includes program code used for performing the methods shown in the flowcharts. In such an embodiment, the computer program may be downloaded and installed from a network through the communication part 1509, and/or installed from the removable medium 1511. When the computer program is executed by the CPU 1501, the various functions defined in the system of this application are executed.

[0131] The computer-readable medium shown in the embodiments of this application may be a computer-readable signal medium or a computer-readable storage medium or any combination of the two. The computer-readable storage medium may be, for example, but is not limited to, an electric, magnetic, optical, electromagnetic, infrared, or semi-conductive system, apparatus, or component, or any combination of the above. A more specific example of the computer-readable storage medium may include but is not limited to: an electrical connection having one or more wires, a portable computer magnetic disk, a hard disk, a random access memory (RAM), a read-only memory (ROM), an erasable programmable read-only memory (EPROM), a flash memory, an optical fiber, a compact disk read-only memory (CD-ROM), an optical storage device, a magnetic storage device, or any appropriate combination thereof. In this application, the computer-readable storage medium may be any tangible medium containing or storing a program, and the program may be used by or used in combination with an instruction execution system, an apparatus, or a device. In this application, a computer-readable signal medium may include a data signal being in a baseband or propagated as a part of a carrier wave, the data signal carrying computer-readable program code. A data signal propagated in such a way may assume a plurality of forms, including, but not limited to, an electromagnetic signal, an optical signal, or any appropriate combination thereof. The computer-readable signal medium may be further any computer readable medium in addition to a computer-readable storage medium. The computer readable medium may send, propagate, or transmit a program that is used by or used in conjunction with an instruction execution system, an apparatus, or a device. The program code included in the computer-readable medium may be transmitted by using any suitable medium, including but not limited to: a wireless medium, a wire, or the like, or any suitable combination thereof.

[0132] The flowcharts and block diagrams in the accompanying drawings illustrate possible system architectures, functions and operations that may be implemented by a system, a method, and a computer program product according to various embodiments of this application. In this regard, each box in a flowchart or a block diagram may represent a module, a program segment, or a part of code. The module, the program segment, or the part of code includes one or more executable instructions used for implementing designated logic functions. In some implementations used as substitutes, functions annotated in boxes may alternatively occur in a sequence different from that annotated in an accompanying drawing. For example, actually two boxes shown in succession may be performed basically in parallel, and sometimes the two boxes may be performed in a reverse sequence. This is determined by a related function. In addition, each box in a block diagram and/or a flowchart and a combination of boxes in the block diagram and/or the flowchart may be implemented by using a dedicated hardware-based system configured to perform a specified function or operation, or may be implemented by using a combination of dedicated hardware and a computer instruction.

[0133] A related unit described in the embodiments of this application may be implemented in a software manner, or may be implemented in a hardware manner, and the unit described can also be set in a processor. Names of the units do not constitute a limitation on the units in a specific case.

[0134] According to another aspect, this application further provides a computer-readable medium. The computer-readable medium may be included in the electronic device described in the foregoing embodiments, or may exist alone and is not disposed in the electronic device. The computer-readable medium carries one or more programs, the one or more programs, when executed by the electronic device, causing the electronic device to implement the method described in the foregoing embodiments.

[0135] According to another aspect, this application further provides a computer program product including instructions, the instructions, when run on a server, causing the server to perform the method in the foregoing embodiments.

[0136] Although a plurality of modules or units of a de-

vice configured to perform actions are discussed in the foregoing detailed description, such division is not mandatory. Actually, according to the implementations of this application, the features and functions of two or more modules or units described above may be specifically implemented in one module or unit. On the contrary, the features and functions of one module or unit described above may be further divided to be embodied by a plurality of modules or units.

[0137] According to the foregoing descriptions of the implementations, a person skilled in the art may readily understand that the exemplary implementations described herein may be implemented by using software, or may be implemented by combining software and necessary hardware. Therefore, the technical solutions of the implementations of this application may be implemented in a form of a software product. The software product may be stored in a non-volatile storage medium (which may be a CD-ROM, a USB flash drive, a removable hard disk, or the like) or on a network, and includes several instructions for instructing a computing device (which may be a personal computer, a server, a touch terminal, network device, or the like) to perform the methods according to the implementations of this application.

[0138] After considering the specification and practicing the disclosed embodiments, a person skilled in the art may easily conceive of other implementations of this application. This application is intended to cover any variations, uses, or adaptive changes of this application. These variations, uses, or adaptive changes follow the general principles of this application and include common general knowledge or common technical means in the art, which are not disclosed in this application.

[0139] It is to be understood that this application is not limited to the precise structures described above and shown in the accompanying drawings, and various modifications and changes can be made without departing from the scope of this application. The scope of this application is subject only to the appended claims.

Claims

1. A communication method, applicable to a core network (CN) entity, the method comprising:

   obtaining a total latency tolerance of uplink (UL) transmission and downlink (DL) transmission of a target service;
   generating a transmission latency indication message for the target service according to the total latency tolerance, the transmission latency indication message comprising quality of service (QoS) flow indication information of the UL transmission of the target service and QoS flow indication information of the DL transmission of the target service; and
   configuring the transmission latency indication message for an access network entity, so as to enable the access network entity to monitor a transmission latency of the target service according to the transmission latency indication message.

2. The communication method according to claim 1, wherein the obtaining a total latency tolerance of UL transmission and DL transmission of a target service comprises:

   receiving a notification message transmitted by an application function (AF) entity, and obtaining the total latency tolerance according to the notification message; or
   obtaining the total latency tolerance based on information of a contractual agreement with the AF entity; or
   obtaining the total latency tolerance pre-configured by the AF entity.

3. The communication method according to claim 1, wherein the generating a transmission latency indication message for the target service according to the total latency tolerance comprises:

   generating information indicating the total latency tolerance according to the total latency tolerance, and allocating QoS flow indication information respectively to the UL transmission and the DL transmission of the target service; and
   generating the transmission latency indication message according to the information indicating the total latency tolerance and the QoS flow indication information respectively allocated to the UL transmission and the DL transmission of the target service.

4. The communication method according to claim 1, wherein the generating a transmission latency indication message for the target service according to the total latency tolerance comprises:

   determining an UL latency tolerance and a DL latency tolerance of the target service according to the total latency tolerance; and
   generating the QoS flow indication information of the UL transmission according to the UL latency tolerance, and generating the QoS flow indication information of the DL transmission according to the DL latency tolerance.

5. The communication method according to claim 4, wherein the determining an UL latency tolerance and a DL latency tolerance of the target service according to the total latency tolerance comprises:

   dividing the total latency tolerance to obtain the

UL latency tolerance and the DL latency tolerance of the target service; or
receiving a division status, notified by an AF entity, of the UL latency tolerance and the DL latency tolerance of the target service, and determining the UL latency tolerance and the DL latency tolerance according to the division status and the total latency tolerance; or
receiving priorities, notified by the AF entity, of the UL transmission and the DL transmission of the target service, and determining the UL latency tolerance and the DL latency tolerance according to the priorities and the total latency tolerance.

6. The communication method according to claim 1, wherein the QoS flow indication information of the UL transmission of the target service comprises:

one piece of QoS flow indication information corresponding to all UL data of the target service; or QoS flow indication information respectively corresponding to different types of UL data of the target service, the different types of UL data being respectively corresponding to different QoS flow indication information.

7. The communication method according to any one of claims 1 to 6, wherein the QoS flow indication information of the DL transmission of the target service comprises:

one piece of QoS flow indication information corresponding to all DL data of the target service; or QoS flow indication information respectively corresponding to different types of DL data of the target service, the different types of DL data being respectively corresponding to different QoS flow indication information.

8. The communication method according to any one of claims 1 to 6, wherein the obtaining a total latency tolerance of UL transmission and DL transmission of a target service comprises:

obtaining, by a policy control function (PCF) entity, the total latency tolerance from the AF entity; or
obtaining, by a network exposure function (NEF) entity, the total latency tolerance from the AF entity and forwarding the total latency tolerance to the PCF entity.

9. A communication method, applicable to an access network entity, the method comprising:

obtaining a transmission latency indication message configured by a core network (CN) entity for a target service, the transmission latency indication message being generated by the CN entity according to a total latency tolerance of uplink (UL) transmission and downlink (DL) transmission of the target service, and the transmission latency indication message comprising quality of service (QoS) flow indication information of the UL transmission of the target service and QoS flow indication information of the DL transmission of the target service; and
monitoring a transmission latency of the target service according to the transmission latency indication message.

10. The communication method according to claim 9, wherein the monitoring a transmission latency of the target service according to the transmission latency indication message comprises:

monitoring a total latency of the UL transmission and the DL transmission of the target service according to the total latency tolerance in a case that the transmission latency indication message further comprises information indicating the total latency tolerance;
monitoring an UL transmission latency of the target service according to an UL latency tolerance in a case that the QoS flow indication information of the UL transmission comprises the UL latency tolerance; and
monitoring a DL transmission latency of the target service according to a DL latency tolerance in a case that the QoS flow indication information of the DL transmission comprises the DL latency tolerance.

11. The communication method according to claim 9, wherein the QoS flow indication information of the UL transmission comprises QoS flow indication information respectively corresponding to different types of UL data of the target service, and the QoS flow indication information of the DL transmission comprises QoS flow indication information respectively corresponding to different types of DL data of the target service; and

the monitoring a transmission latency of the target service according to the transmission latency indication message comprises: monitoring, according to the transmission latency indication message, a transmission latency of UL data of a specified type of the target service and a transmission latency of DL data of a specified type of the target service,
the UL data of the specified type and the DL data of the specified type being located in the same protocol data unit (PDU) session or in different PDU sessions during transmission.

12. The communication method according to claim 11, wherein the target service comprises a cloud gaming service; and
the UL data of the specified type comprises game manipulation data, and the DL data of the specified type comprises multimedia data obtained by rendering a game scene.

13. The communication method according to any one of claims 9 to 12, wherein the monitoring a transmission latency of the target service according to the transmission latency indication message comprises:

monitoring a transmission latency of the target service between user equipment (UE) and the access network entity according to the transmission latency indication message; and/or
monitoring a transmission latency of the target service between UE and a user plane function (UPF) entity according to the transmission latency indication message and latency information between the access network entity and the UPF entity.

14. A communication apparatus, comprising:

an obtaining unit, configured to obtain a total latency tolerance of uplink (UL) transmission and downlink (DL) transmission of a target service;
a generation unit, configured to generate a transmission latency indication message for the target service according to the total latency tolerance, the transmission latency indication message comprising quality of service (QoS) flow indication information of the UL transmission of the target service and QoS flow indication information of the DL transmission of the target service; and
a configuration unit, configured to configure the transmission latency indication message for a access network entity, so as to enable the access network entity to monitor a transmission latency of the target service according to the transmission latency indication message.

15. A communication apparatus, comprising:

an obtaining unit, configured to obtain a transmission latency indication message configured by a core network (CN) entity for a target service, the transmission latency indication message being generated by the CN entity according to a total latency tolerance of uplink (UL) transmission and downlink (DL) transmission of the target service, and the transmission latency indication message comprising quality of service (QoS) flow indication information of the UL transmission of the target service and QoS flow indication information of the DL transmission of the target service; and
a monitoring unit, configured to monitor a transmission latency of the target service according to the transmission latency indication message.

16. A computer-readable medium, storing a computer program, the computer program, when executed by a processor, implementing the communication method according to any one of claims 1 to 8, or implementing the communication method according to any one of claims 9 to 13.

17. An electronic device, comprising:

one or more processors; and
a storage apparatus, configured to store one or more programs, the one or more programs, when executed by the one or more processors, causing the one or more processors to implement the communication method according to any one of claims 1 to 8, or implement the communication method according to any one of claims 9 to 13.

18. A computer program product comprising instructions, the instructions, when run on a computer, causing the computer to perform the communication method according to any one of claims 1 to 8, or perform the communication method according to any one of claims 9 to 13.

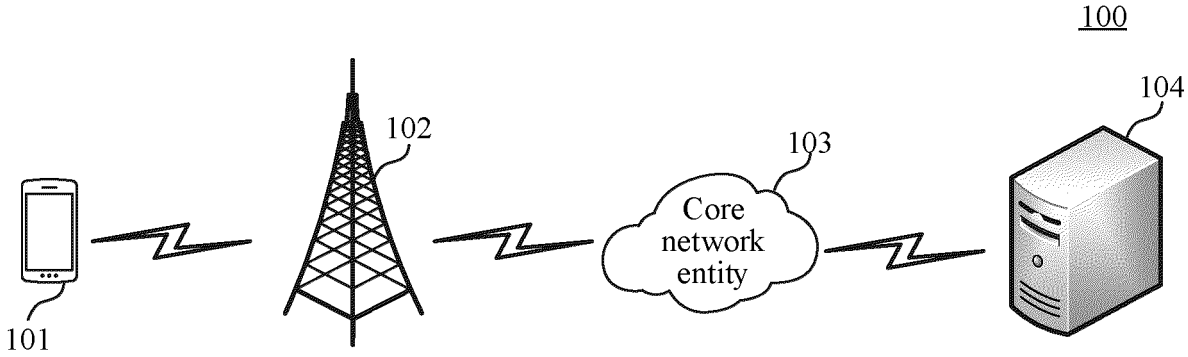

FIG. 1

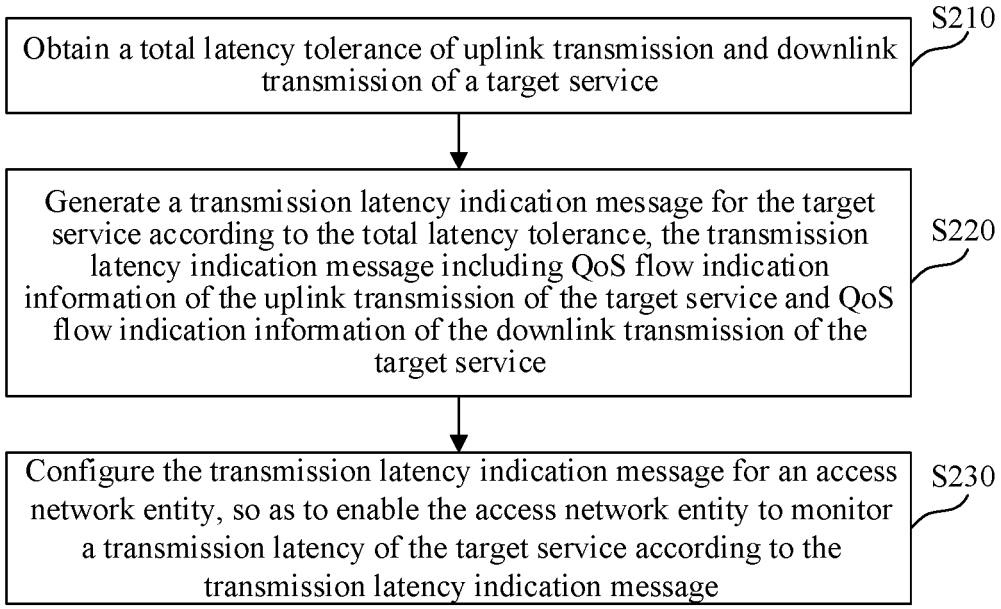

FIG. 2

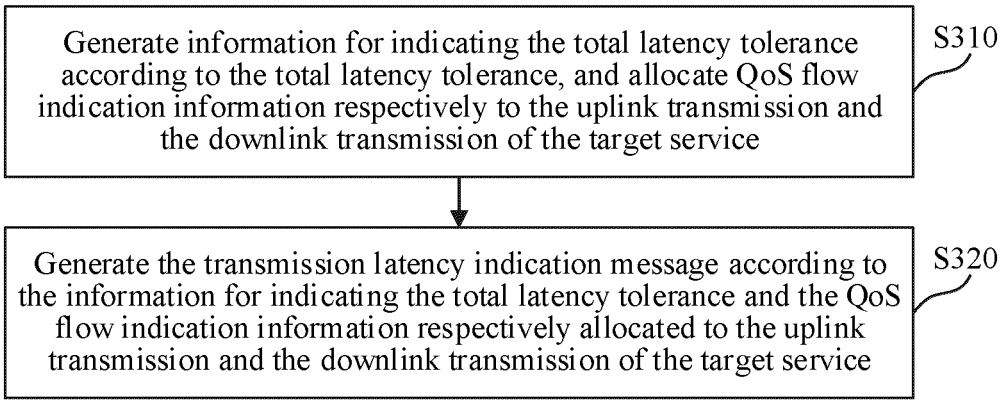

FIG. 3

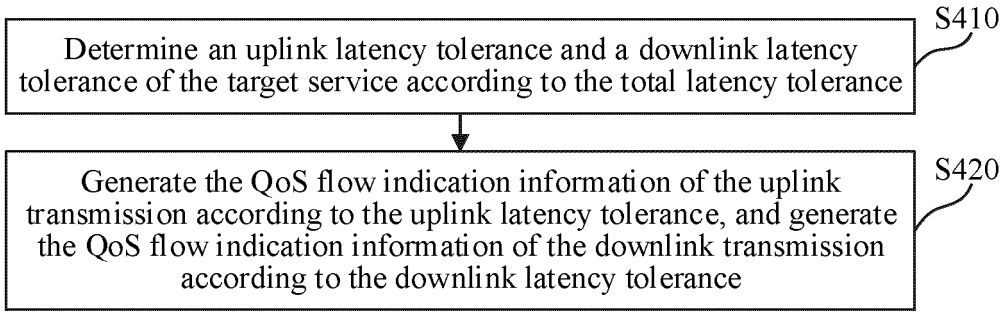

FIG. 4

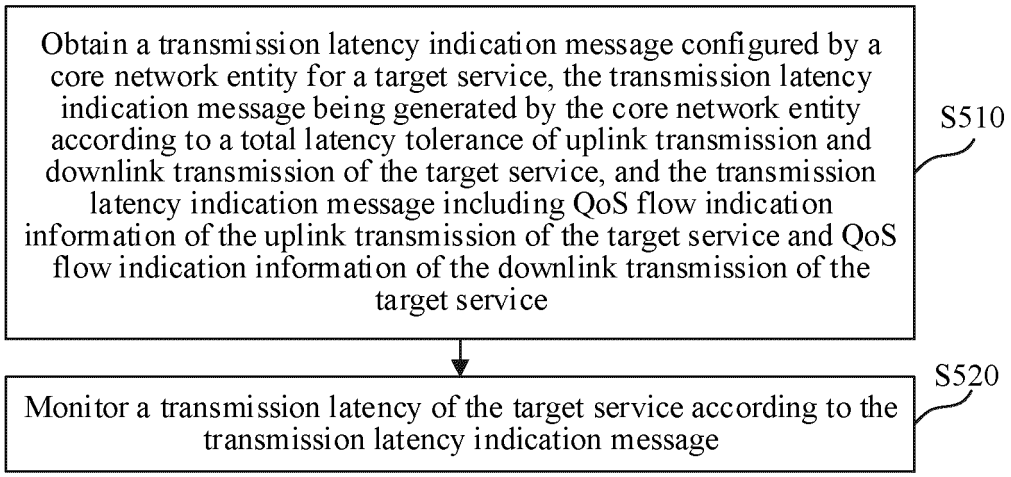

FIG. 5

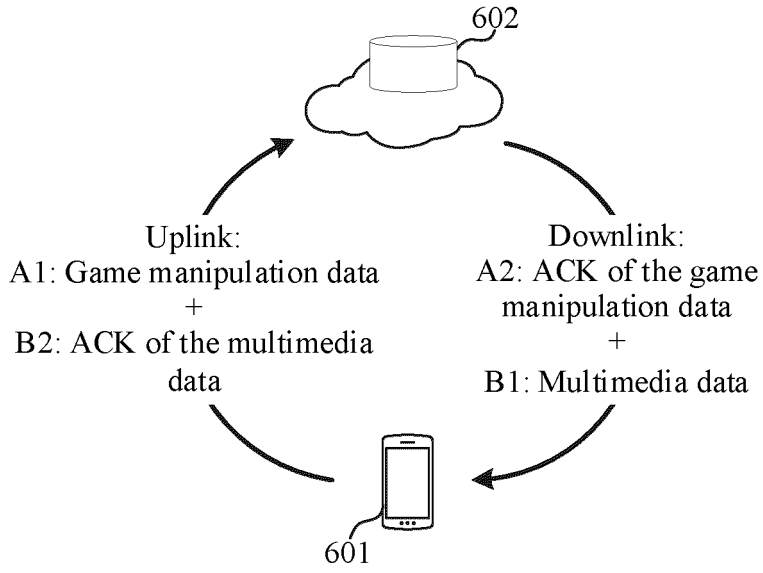

FIG. 6

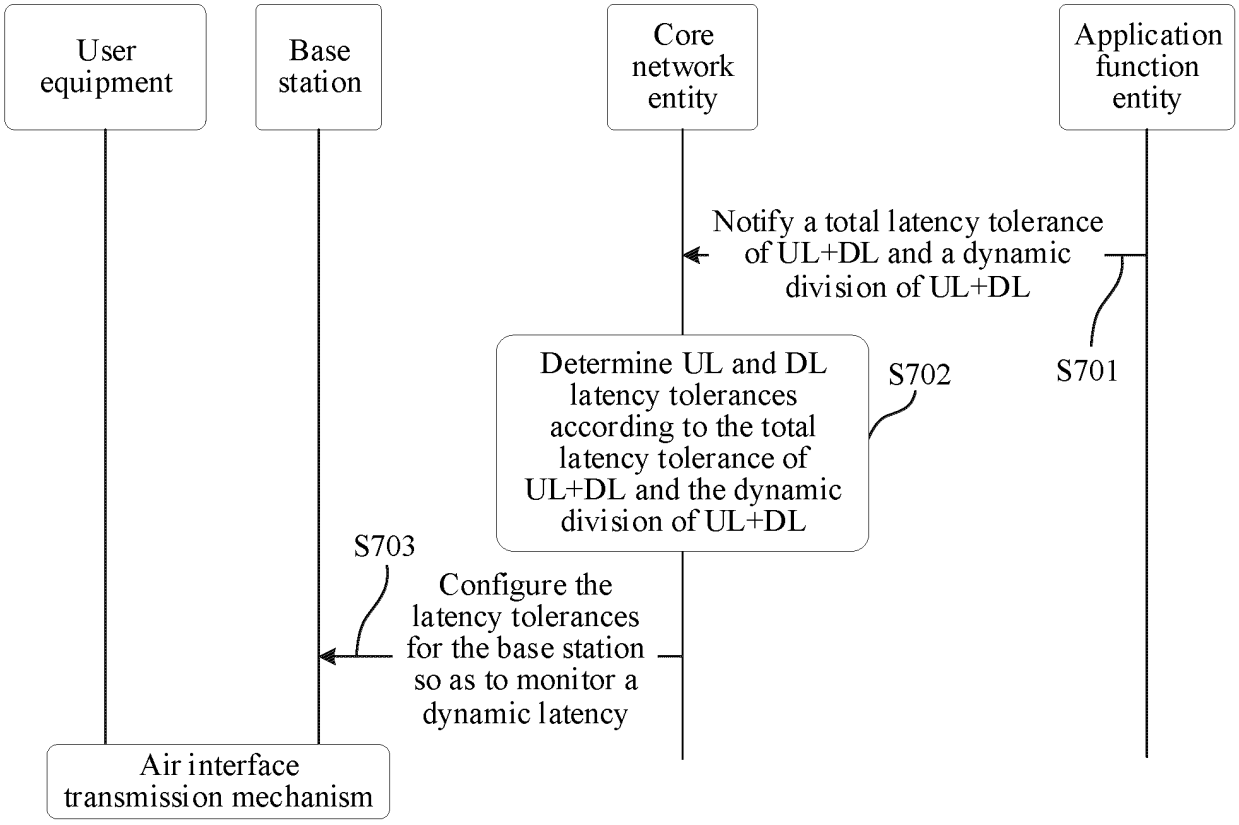
User equipment
Base station
Core network entity
Application function entity
Notify a total latency tolerance of UL+DL and a dynamic division of UL+DL
S701
Determine UL and DL latency tolerances according to the total latency tolerance of UL+DL and the dynamic division of UL+DL
S702
S703
Configure the latency tolerances for the base station so as to monitor a dynamic latency
Air interface transmission mechanism

FIG. 7

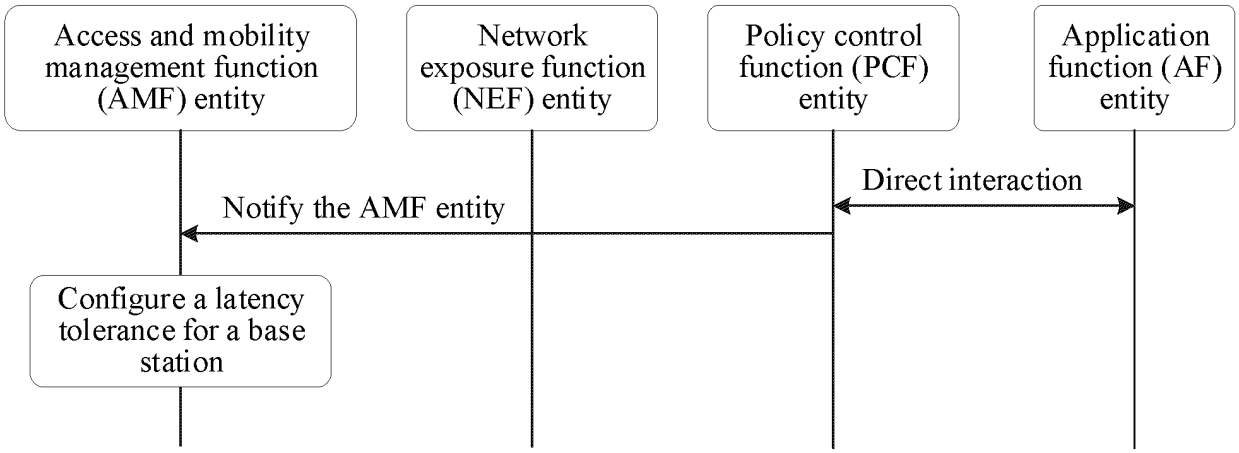
Access and mobility management function (AMF) entity
Network exposure function (NEF) entity
Policy control function (PCF) entity
Application function (AF) entity
Direct interaction
Notify the AMF entity
Configure a latency tolerance for a base station

FIG. 8

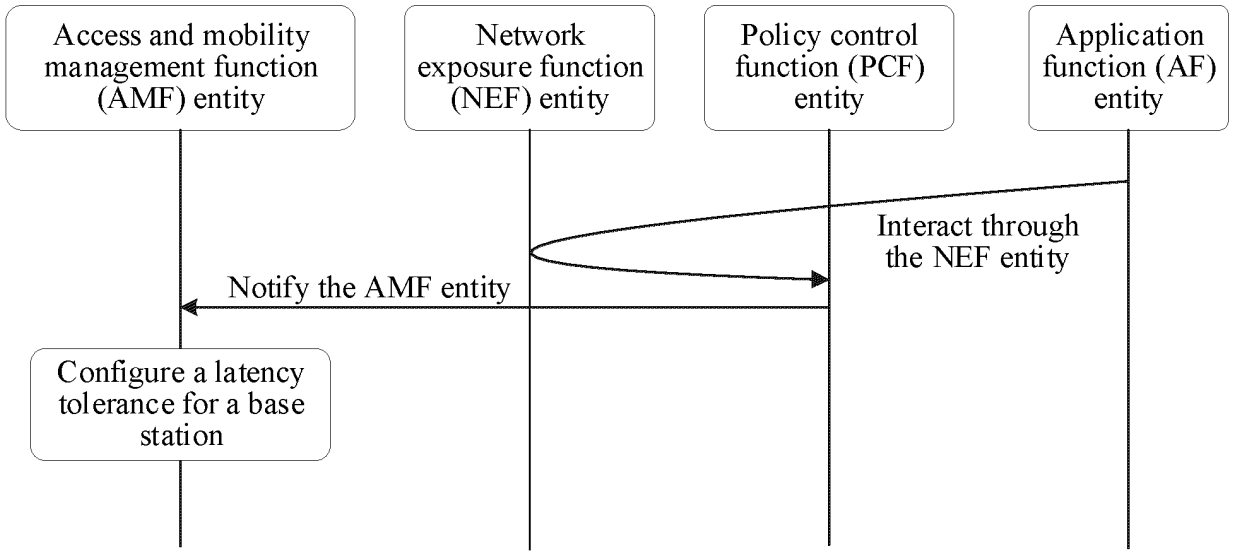
Access and mobility management function (AMF) entity
Network exposure function (NEF) entity
Policy control function (PCF) entity
Application function (AF) entity
Interact through the NEF entity
Notify the AMF entity
Configure a latency tolerance for a base station

FIG. 9

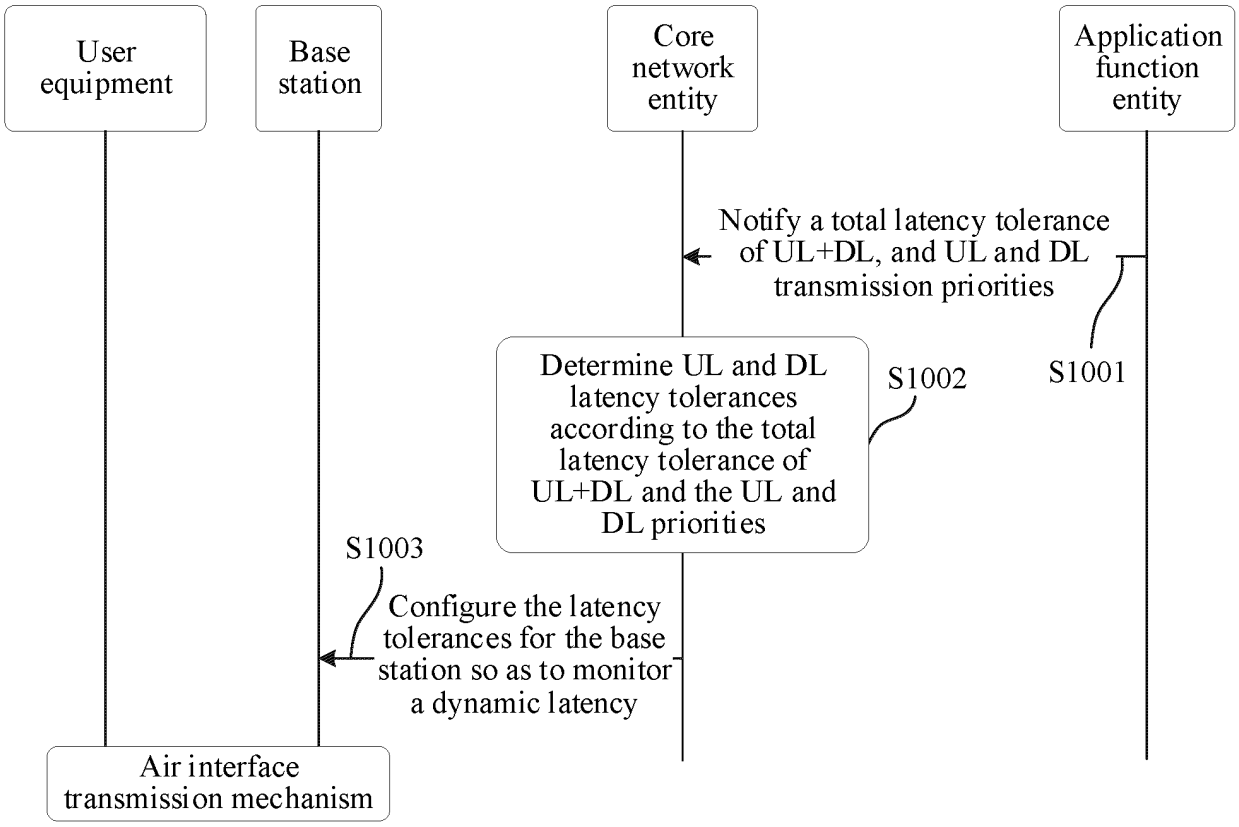
User equipment
Base station
Core network entity
Application function entity
Notify a total latency tolerance of UL+DL, and UL and DL transmission priorities
S1001
Determine UL and DL latency tolerances according to the total latency tolerance of UL+DL and the UL and DL priorities
S1002
S1003
Configure the latency tolerances for the base station so as to monitor a dynamic latency
Air interface transmission mechanism

FIG. 10

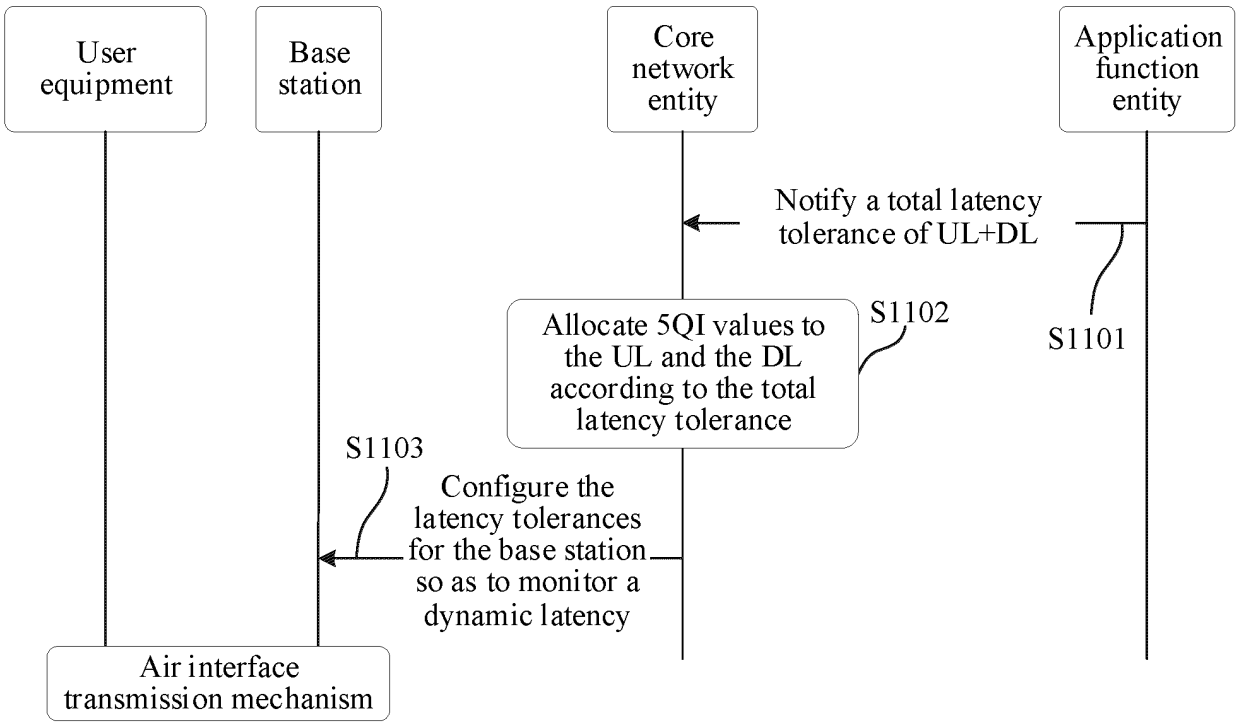
User equipment
Base station
Core network entity
Application function entity
Notify a total latency tolerance of UL+DL
S1101
Allocate 5QI values to the UL and the DL according to the total latency tolerance
S1102
S1103
Configure the latency tolerances for the base station so as to monitor a dynamic latency
Air interface transmission mechanism

FIG. 11

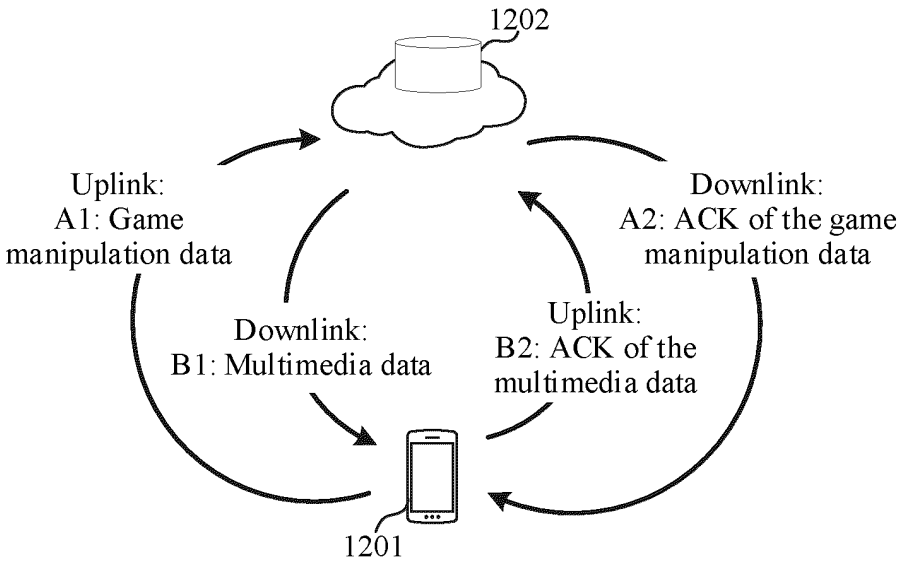
1202
Uplink:
A1: Game manipulation data
Downlink:
A2: ACK of the game manipulation data
Downlink:
B1: Multimedia data
Uplink:
B2: ACK of the multimedia data
1201

FIG. 12

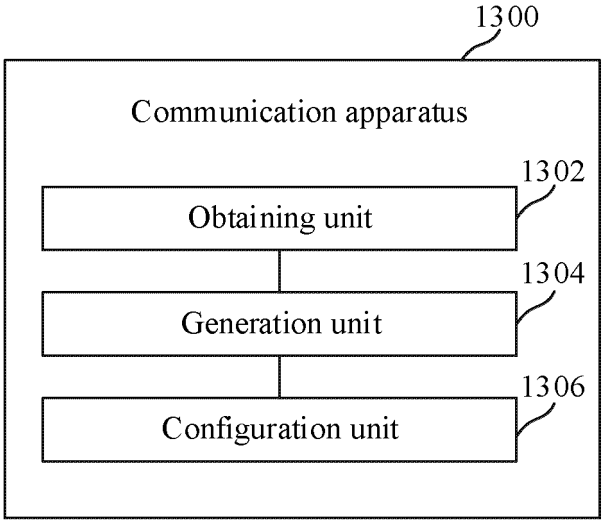
1300
Communication apparatus
1302
Obtaining unit
1304
Generation unit
1306
Configuration unit

FIG. 13

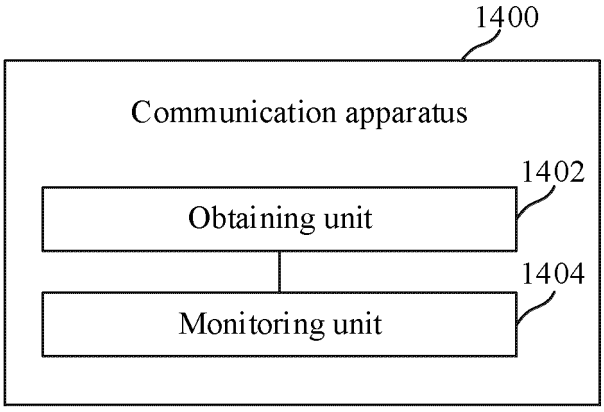
1400
Communication apparatus
1402
Obtaining unit
1404
Monitoring unit

FIG. 14

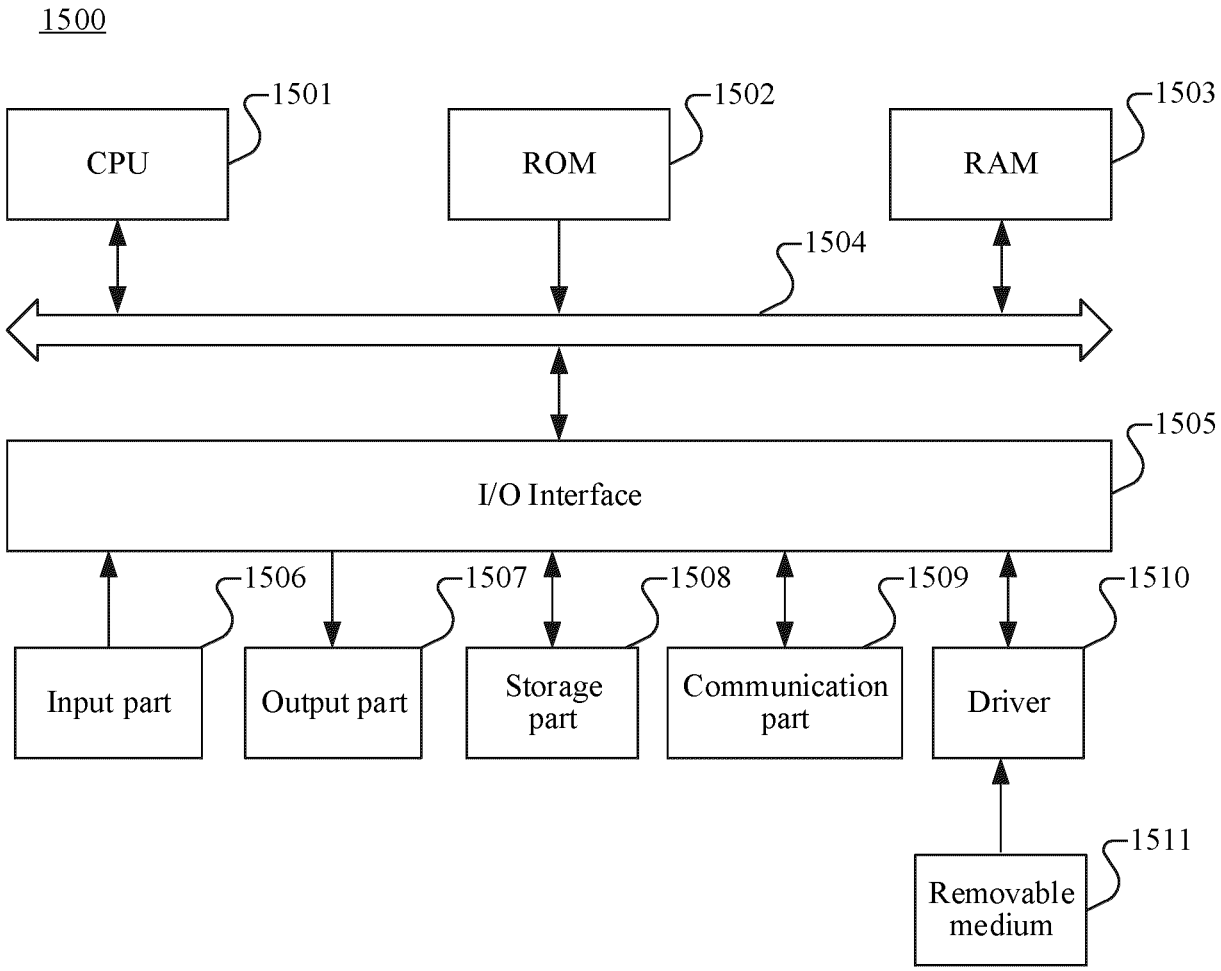
1500
1501
CPU
1502
ROM
1503
RAM
1504
1505
I/O Interface
1506
Input part
1507
Output part
1508
Storage part
1509
Communication part
1510
Driver
1511
Removable medium

FIG. 15

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2020/076623

A. CLASSIFICATION OF SUBJECT MATTER

H04W 28/24(2009.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

H04W H04L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNKI, CNABS, CNTXT, VEN, USTXT, WOTXT, EPTXT: 上行, 上, 下行, 时延, 延时, 延迟, 游戏, 容限, 上限, 需求, 要求, 最大值, 限制值, 阈值, 容忍, 监测, 监控, 总, 和, 基站, 接入网, 核心网, 实体, 指示, 分割, 分配; UL, DL, uplink, downlink, delay, threshold, requirement, total, QoS, segment, distribution, allocation, game

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 110062426 A (TENCENT TECHNOLOGY SHENZHEN CO., LTD.) 26 July 2019 (2019-07-26) claims 1-17, and description, paragraph [0159]	1-18
A	CN 109245936 A (HUAWEI TECHNOLOGIES CO., LTD.) 18 January 2019 (2019-01-18) description, paragraphs [0053]-[0240]	1-18
A	CN 102111819 A (TSINGHUA UNIVERSITY) 29 June 2011 (2011-06-29) entire document	1-18
A	CN 102638852 A (CHINA ACADEMY OF TELECOMMUNICATIONS TECHNOLOGY) 15 August 2012 (2012-08-15) entire document	1-18
A	WO 2018130741 A1 (NOKIA TECHNOLOGIES OY) 19 July 2018 (2018-07-19) entire document	1-18
A	WO 2018145103 A1 (IDAC HOLDINGS INC.) 09 August 2018 (2018-08-09) entire document	1-18

☐ Further documents are listed in the continuation of Box C. ☑ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
21 May 2020	**29 May 2020**
Name and mailing address of the ISA/CN **China National Intellectual Property Administration (ISA/CN)** **No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088** **China** Facsimile No. **(86-10)62019451**	Authorized officer Telephone No.

Form PCT/ISA/210 (second sheet) (January 2015)

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2020/076623

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
CN	110062426	A	26 July 2019	None			
CN	109245936	A	18 January 2019	EP	2892190	A1	08 July 2015
				EP	2892190	A4	02 September 2015
				US	2015189539	A1	02 July 2015
				CN	103782555	A	07 May 2014
				US	9936416	B2	03 April 2018
				WO	2014036704	A1	13 March 2014
				IN	201501922	P4	01 July 2016
CN	102111819	A	29 June 2011	CN	102111819	B	01 January 2014
CN	102638852	A	15 August 2012	EP	2675234	A4	12 March 2014
				US	8842629	B2	23 September 2014
				CN	102638852	B	22 June 2016
				EP	2675234	B1	21 September 2016
				US	2014056245	A1	27 February 2014
				WO	2012107004	A1	16 August 2012
				EP	2675234	A1	18 December 2013
WO	2018130741	A1	19 July 2018	EP	3569028	A1	20 November 2019
WO	2018145103	A1	09 August 2018	None			

Form PCT/ISA/210 (patent family annex) (January 2015)

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- CN 201910261362 **[0001]**